(FORM PTO-1390 (Modified) (REV 11-2000)) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER	
PB-9944	
U.S. APPLICATION NO (IF KNOWN, SEE 37 CFR	
To be assigned 10/049358	
INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE
PCT/US00/22150	10 August 2000
PRIORITY DATE CLAIMED	
10 August 1999	

TITLE OF INVENTION

TAQ DNA Polymerases Having an Amino Acid Substitution at E681 and Homologs Thereof Exhibiting Improved Salt Tolerance

APPLICANT(S) FOR DO/EO/US

Maria Davis, John Nelson, Shiv Kumar, Patrick Finn, Satyam Nampalli, and Parke Flick

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include itens (5), (6), (9) and (24) indicated below.
4. ☒ The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
   a. ☒ is attached hereto (required only if not communicated by the International Bureau).
   b. ☒ has been communicated by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto.
   b. ☐ has been previously submitted under 35 U.S.C. 154(d)(4).
7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d ☐ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).
11. ☒ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☒ A copy of the International Search Report (PCT/ISA/210).

Items 13 to 20 below concern document(s) or information included:

13. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
15. ☐ A **FIRST** preliminary amendment.
16. ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
17. ☐ A substitute specification.
18. ☐ A change of power of attorney and/or address letter.
19. ☒ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C 1.821 - 1.825
20. ☐ A second copy of the published international application under 35 U.S.C 154(d)(4).
21. ☐ A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
22. ☒ Certificate of Mailing by Express Mail
23. ☒ Other items or information.

additional copy of this transmittal letter for charging purposes
return postcard

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR) To be Assigned 10/049358	INTERNATIONAL APPLICATION NO. PCT/US00/22150	ATTORNEY'S DOCKET NUMBER PB-9944

24. The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

		CALCULATIONS PTO USE ONLY
☐	Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1040.00**	
☒	International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$890.00**	
☐	International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO **$740.00**	
☐	International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) **$710.00**	
☐	International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) **$100.00**	
	ENTER APPROPRIATE BASIC FEE AMOUNT =	**$890.00**
	Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (e)).	**$0.00**

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	
Total claims	28 - 20 =	8	x $18.00	**$144.00**
Independent claims	6 - 3 =	3	x $84.00	**$252.00**
Multiple Dependent Claims (**check if applicable**).			☐	**$0.00**
	TOTAL OF ABOVE CALCULATIONS =			**$1,286.00**
☐	Applicant claims small entity status. See 37 CFR 1.27). The fees indicated above are reduced by 1/2.			**$0.00**
	SUBTOTAL =			**$1,286.00**
	Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (f)). +			**$0.00**
	TOTAL NATIONAL FEE =			**$1,286.00**
☐	Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) (**check if applicable**).			**$0.00**
	TOTAL FEES ENCLOSED =			**$1,286.00**
			Amount to be: refunded	$
			charged	$

a. ☐ A check in the amount of _____ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No. **500-588** in the amount of **$1,286.00** to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. **500-588** A duplicate copy of this sheet is enclosed.

d. ☐ Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Royal N. Ronning, Jr.
Amersham Biosciences Corp.
800 Centennial Avenue
Piscataway, New Jersey 08855

(732) 457-8423

SIGNATURE

Royal N. Ronning, Jr.
NAME

32,529
REGISTRATION NUMBER

February 8, 2002
DATE

PB-9944

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of:	M. Davis, et al.	Group Art Unit:	To be assigned
Serial Number:	10/049,358	Examiner:	To be assigned
Filing Date:	To be assigned		
Title:	*TAQ* DNA Polymerases Having an Amino Acid Substitution at E681 and Homologs Thereof Exhibiting Improved Salt Tolerance		

Submission of Nucleotide and/or Amino Acid Sequence Disclosures

Assistant Commissioner for Patents
Box PCT
Washington, D.C. 20231

Dear Sir:

In connection with the prosecution of the captioned application, and in response to the Notification of Missing Requirements Under 35 U.S.C. 371 in the United States Designated/Elected Office (DO/EO/US) having a mailing date of April 23, 2002, Applicants submit the following items:

1) An initial computer readable form (CFR) copy of the sequence listing and an initial paper copy of the sequence listing, both generated on an IBM computer using PatentIn, Version 2.1 Software.
2) Required statements:
   A. Applicants aver that the sequence listings do not include any new matter which goes beyond the content of the application as filed.

B. Applicants aver that that sequence listings contained on the diskette are identical to those contained on the paper copy.

A copy of the Notification of Missing Requirements Under 35 U.S.C. 371 in the United States Designated/Elected Office (DO/EO/US) having a mailing date of April 23, 2002, is enclosed herewith as required. Please direct any issues to Applicant's counsel at the telephone number provided below.

Respectfully submitted,

Royal N. Ronning, Jr., 32,529
Attorney for Applicants

Amersham Biosciences Corp.
800 Centennial Avenue
P. O. Box 1327
Piscataway, New Jersey 08855-1327

Tel: (732) 457-8423
Fax: (732) 457-8463

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, on May 8, 2002

Melissa Leck
Signature
Date May 8, 2002

H:\DEPT\Law_Dept\IP Dept\Submission of Sequence Listing\Submission of Sequence Listing3-PB9944.doc

PB-9944

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of:	M. Davis, et al.	Group Art Unit:	To be assigned
Serial Number:	To be assigned	Examiner:	To be assigned
Filing Date:	To be assigned		
Title:	*TAQ* DNA Polymerases Having an Amino Acid Substitution at E681 and Homologs Thereof Exhibiting Improved Salt Tolerance		

SUBMISSION OF NUCLEOTIDE AND/OR AMINO ACID SEQUENCE DISCLOSURES

Assistant Commissioner for Patents
Box New Patent Application
Washington, D.C. 20231

Dear Sir:

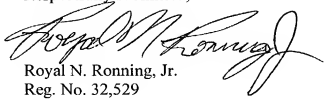
In connection with the prosecution of the captioned application, Applicants submit the following items:

1) An initial computer readable form (CFR) copy of the sequence listing and an initial paper copy of the sequence listing, both generated on an IBM computer using PatentIn, Version 2.1 Software.
2) Required statements:
   A. Applicants aver that the sequence listings do not include any new matter which goes beyond the content of the application as filed.
   B. Applicants aver that that sequence listings contained on the diskette are identical to those contained on the paper copy.

Please direct any issues to Applicant's counsel at the telephone number provided below.

Amersham Biosciences Corp.
P.O. Box 1327
800 Centennial Avenue
Piscataway, New Jersey 08855-1327
(732)457-8423 (voice)
(732)457-8463 (facsimile)

Respectfully submitted,



Royal N. Ronning, Jr.
Reg. No. 32,529

PB-9944

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of:	M. Davis, et al.	Group Art Unit:	To be assigned
Serial Number:	To be assigned	Examiner:	To be assigned
Filing Date:	To be assigned		
Title:	*TAQ* DNA Polymerases Having an Amino Acid Substitution at E681 and Homologs Thereof Exhibiting Improved Salt Tolerance		

First Preliminary Amendment

Honorable Assistant Commissioner of Patents
Box Patent Application
Washington, D.C. 20231

Sir:

Please consider the following amendments and remarks in connection with the prosecution of the captioned application, which is a filing under 35 U.S.C. § 371 and claims priority to international application number PCT/US00/22150 filed August 10, 2000. This application also claims the benefit of United States provisional application number 60/148,012 having a filing date of August 10, 1999.

In the Claims

Please amend claim 1 as follows:

1. (once amended) A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 2 (SEQ ID No. 2).

Please amend claim 2 as follows:

2. (once amended) A purified recombinant thermostable DNA polymerase comprising an amino acid substitution corresponding to the substitution E681R in the amino acid sequence listing of Figure 2 (SEQ ID No. 2).

Please amend claim 3 as follows:

3. (once amended) An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence corresponding to the amino acid sequence set forth in Figure 2 (SEQ ID No. 2).

Please amend claim 15 as follows:

15. (once amended) A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 3 (SEQ ID No. 3).

Please amend claim 16 as follows:

16. (once amended) A DNA polymerase comprising an amino acid substitution corresponding to the substitution E681R in the amino acid sequence listing of Figure 3 (SEQ ID No. 3).

Please amend claim 17 as follows:

17. (once amended) An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence corresponding to the amino acid sequence set forth in Figure 3 (SEQ ID No. 3).

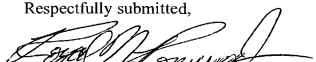
Remarks

Claims 1-28 are pending in the instant application. Applicants have amended claims 1, 2, 3, 15, 16, and 17 to more fully conform with U.S. practice. A version of the claims marked up to show the amendments, as well as a clean version of the claims encompassing the amendments, is attached hereto.

Applicants respectfully assert that all amendments are fairly based on the specification, and respectfully request their entry.

Applicants believe that the claims, as amended, are in allowable form, and earnestly solicit the allowance of claims 1-28.

Respectfully submitted,



Royal N. Ronning, Jr. 32,529
Attorney for Applicants

Amersham Biosciences
800 Centennial Avenue
P. O. Box 1327
Piscataway, New Jersey 08855-1327

Tel: (732) 457-8423
Fax: (732) 457-8463

Claims (marked up version showing amendments)

1. (once amended) A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 2 (SEQ ID No. 2).

2. (once amended) A purified recombinant thermostable DNA polymerase comprising an amino acid substitution corresponding to the substitution E681R in the amino acid sequence listing of Figure 2 (SEQ ID No. 2).

3. (once amended) An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence corresponding to the amino acid sequence set forth in Figure 2 (SEQ ID No. 2).

15. (once amended) A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 3 (SEQ ID No. 3).

16. (once amended) A DNA polymerase comprising an amino acid substitution corresponding to the substitution E681R in the amino acid sequence listing of Figure 3 (SEQ ID No. 3).

17. (once amended) An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence corresponding to the amino acid sequence set forth in Figure 3 (SEQ ID No. 3).

Claims (clean version encompassing amendments)

1. (once amended) A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 2 (SEQ ID No. 2).

2. (once amended) A purified recombinant thermostable DNA polymerase comprising an amino acid substitution corresponding to the substitution E681R in the amino acid sequence listing of Figure 2 (SEQ ID No. 2).

3. (once amended) An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence corresponding to the amino acid sequence set forth in Figure 2 (SEQ ID No. 2).

4. A recombinant DNA vector that comprises the nucleic acid of Claim 3.

5. A recombinant host cell transformed with the vector of Claim 4.

6. The recombinant host cell of Claim 5 that is *E. coli.*

7. A method of sequencing DNA comprising the step of generating chain terminated fragments from the DNA template to be sequenced with the DNA polymerase of Claim 1 in the presence of at least one chain terminating agent and one or more

nucleotide triphosphates, and determining the sequence of said DNA from the sizes of said fragments.

8. A method according to Claim 7, wherein the chain terminating agent comprises a labeled nucleic acid terminator having a net positive or a net negative charge.

9. A method for synthesizing a fluorescently labeled polynucleotide, said method comprising the step of mixing a DNA polymerase according to Claim 1 with a primed template.

10. A method according to Claim 9, wherein the primed template is a primed template in a chain termination sequencing reaction.

11. A method according to Claim 9, wherein the primed template is a primed template in a polymerase chain reaction.

12. A kit for fluorescently labeling a polynucleotide, the kit comprising a DNA polymerase according to Claim 1 and a fluorescently labeled nucleotide.

13. A kit according to Claim 12, wherein the fluorescently labeled nucleotide comprises a nucleic acid terminator having a net negative or a net positive charge.

14. A kit for sequencing DNA comprising the DNA polymerase of Claim 1.

15. (once amended) A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 3 (SEQ ID No. 3).

16. (once amended) A DNA polymerase comprising an amino acid substitution corresponding to the substitution E681R in the amino acid sequence listing of Figure 3 (SEQ ID No. 3).

17. (once amended) An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence corresponding to the amino acid sequence set forth in Figure 3 (SEQ ID No. 3).

18. A recombinant DNA vector that comprises the nucleic acid of Claim 17.

19. A recombinant host cell transformed with the vector of Claim 18.

20. The recombinant host cell of Claim 18 that is *E. coli.*

21. A method of sequencing DNA comprising the step of generating chain terminated fragments from the DNA template to be sequenced with the DNA polymerase of Claim 16 in the presence of at least one chain terminating agent and one or more nucleotide triphosphates, and determining the sequence of said DNA from the sizes of said fragments.

22. A method according to Claim 21, wherein the chain terminating agent comprises a labeled nucleic acid terminator having a net positive or a net negative charge.

23. A method for synthesizing a fluorescently labeled polynucleotide, said method comprising the step of mixing a DNA polymerase according to Claim 16 with a primed template.

24. A method according to Claim 23, wherein the primed template is a primed template in a chain termination sequencing reaction.

25. A method according to claim 23, wherein the primed template is a primed template in a polymerase chain reaction.

26. A kit for fluorescently labeling a polynucleotide, the kit comprising a DNA polymerase according to Claim 16 and a fluorescently labeled nucleotide.

27. A kit according to Claim 26, wherein the fluorescently labeled nucleotide comprises a nucleic acid terminator having a net negative or a net positive charge.

28. A kit for sequencing DNA comprising the DNA polymerase of Claim 16.

TAQ DNA POLYMERASES HAVING AN AMINO ACID SUBSTITUTION AT E681 AND HOMOLOGS THEREOF EXHIBITING IMPROVED SALT TOLERANCE

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority under 35 USC § 119(e) of US provisional application serial number 60/150,167, filed on August 21, 1999, and US provisional application serial number 60/154,739, filed on September 17, 1999, the entire disclosures of each of which are incorporated in their entirety herein.

BACKGROUND OF THE INVENTION

Field of the Invention

The instant disclosure pertains to thermostable DNA polymerases which exhibit improved robustness and efficiency. In particular, the instant DNA polymerase has been shown to result in a substantial improvement of signal uniformity compared to Taq Δ271/ F272M/F667Y DNA polymerase when used in DNA sequencing reactions.

Background

DNA polymerases are enzymes which are useful in many recombinant DNA techniques such as nucleic acid amplification by the polymerase chain reaction ("PCR"), self-sustained sequence replication ("3SR"), and high temperature DNA sequencing. Thermostable polymerases are particularly useful. Because heat does not destroy the polymerase activity, there is no need to add additional polymerase after every denaturation step.

Naturally occurring DNA polymerases preferentially incorporate unlabeled nucleotides over corresponding labeled nucleotides into polynucleotides. This ability of DNA polymerases to discriminate against fluorescently labeled nucleotides had an undesirable effect on many molecular biology procedures that require the enzymatic addition of labeled nucleotides, e.g., labeled dideoxy terminator sequencing. Ambiguous sequencing determinations often result from the disproportionate number of labeled and unlabeled dideoxy terminators and nucleotides. On an electropherogram obtained from a capillary

electrophoresis sequencing unit, this phenomena shows up as uneven peaks. Large signals due to a larger amount of incorporated labeled ddNTP (shown as wide peaks) can obscure smaller signals and lead to ambiguous sequence determinations. Additionally, many of the enzymes presently available are sensitive to high salt environments.

Thus, a need continues to exist for an improved DNA polymerase having improved discrimination properties (and thus resulting in improved signal uniformity) and increased tolerance to high salt conditions. These and other concerns are addressed in greater detail below.

BRIEF SUMMARY OF THE INVENTION

The instant disclosure teaches a purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 2 or 3. The instant disclosure also teaches an isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence set forth in Figure 2 or 3, as well as a recombinant DNA vector that comprises the nucleic acid, and a recombinant host cell transformed with the vector. The instant disclosure also teaches a method of sequencing DNA comprising the step of generating chain terminated fragments from the DNA template to be sequenced with the DNA polymerase in the presence of at least one chain terminating agent having a net negative or a net positive charge and one or more nucleotide triphosphates, and determining the sequence of said DNA from the sizes of said fragments. The instant disclosure also teaches a kit for sequencing DNA comprising the DNA polymerase and nucleic acid terminator having a net negative or a net positive charge.

DETAILED DESCRIPTION

One objective of the instant disclosure is to increase the uniformity of dye-terminator incorporation in fluorescent dye DNA sequencing. One important DNA polymerase is Taq DNA polymerase isolated from the thermophilic bacterium *Thermus aquaticus*, the amino acid sequence for which is shown at Figure 1. The full length enzyme was truncated to eliminate 5' to 3' exonuclease activity and to provide a polypeptide more stable to proteolysis and heat treatment. The truncated enzyme is known as Taq Δ271/ F272M/F667Y DNA polymerase, which is commercially available from Amersham Pharmacia Biotech as Thermo Sequenase® DNA polymerase. Position 1 (amino acid Met) in Taq Δ271/ F272M/F667Y

DNA polymerase corresponds to position 272 in full length Taq polymerase. It should be noted that the numbering used in the instant disclosure is that for Taq Δ271/ F272M/F667Y polymerase, not for Taq polymerase.

Single amino acid substitutions were introduced into Taq Δ271/ F272M/F667Y polymerase. These substitutions are designated as E344Q, I367V, F367Y, E416K and E410R. Each of the subsituted polymerases was expressed, purified, and analyzed for uniformity of dye-terminator incorporation in fluorescent sequencing studies, as assayed by signal uniformity. The E410R substitution was found to result in a substantial improvement of signal uniformity compared to Taq Δ271/ F272M/F667Y DNA polymerase.

The DNA polymerases disclosed herein are especially suitable for use in sequencing reactions which employ terminators having a net positive or a net negative charge. Surprisingly, the instant DNA polymerases have been shown to modulate the incorporation of such terminators during the sequencing reaction. See for example Figure 14. Furthermore, such nucleic acid terminators, which along with the corresponding nucleic acid terminator decomposition products, migrate on separation media at different rates than the sequencing reaction products and which result in improved sequence data. These nucleic acid terminators also allow for the direct loading of nucleic acid sequencing reactions onto separating media. To achieve this goal, negatively or positively charged moieties are attached to the terminator molecule. The unreacted or degraded terminators containing such charged moieties move faster (negatively charged) or in the reverse direction (positively charged) than the DNA sequencing products.

For example, the structures depicted in Figure 15 illustrate potential sites at which a charged moiety may be attached to a terminator. Referring to Figure 15, the Base may comprise A, T, G, C or analogs such as 7-deazapurine, inosine, universal bases. The Sugar may comprise furanose, hexose, mono-di-triphosphates, morpholine, didehydro, dideoxyribose, deoxyribose. The Linker may comprise 1-100 atoms, preferably 2-50 atoms consisting of C, H, N, O, S and halogens. The Mobility modifier may comprise any charged species which alters electrophoretic mobility of structure and degradation products, e.g., α-sulfo-β-alanine, cysteic acid, sulfonic acids, carboxylates, phosphates, phosphodiesters, phosphonates, amines, quaternised amines, and phosphonium moieties. The Mobility modifier may comprise a number of these units covalently linked together. The Label may comprise any signal moiety such as radioisotope, electrochemical tag, fluorescent tags,

energy transfer (ET) labels, mass spectrometry tags, Raman tags, hapten, chemilluminescent group, enzyme, chromophore, and two or more labels. The label may also be charged, e.g. Cy5.5, bis-sulfonated carboxyfluorescein, or a dye attached to a charged moiety, e.g., carboxyfluorescein attached to cysteic acid or similar charged species. Methods for making these and other compounds are disclosed in U.S. Provisional Application No. 60/098,469 filed on August 31, 1998, and U.S. Application No. 90/018,695 filed on February 4, 1998, and PCT/GB98/00978 filed on April 2, 1998 and published on October 8, 1998, the disclosures of each application are incorporated in their entirety by reference herein.

The following examples are for illustration purposes only and should not be used in any way to limit the appended claims.

EXAMPLES

EXAMPLE 1

The construction, expression and purification of Taq Δ271/F272M/F667Y/E410R polymerase is described below. The other substitutions named above were constructed, expressed and purified in a similar manner.

Construction

Primers BamHIFOR (5' ccg ctt ggg cag agg atc cgc cgg gcc ttc atc gcc gag ga) and NheIREV (5' tcg taa ggg atg gct agc cgc tgg gag agg cgg tgg gcc gac) were used in a standard PCR reaction to amplify the region between the BamHI and NheI restriction sites in pREFY2pref (cloned Taq Δ271/F272M/F667Y DNA polymerase). Primer BamHIFOR contains a BamHI restriction site which corresponds to the same unique site in pREFY2pref, and primer NheIREV contains a NheI restriction site which corresponds to the same unique site in pREFY2pref. In addition, primer NheIREV was designed to change the codon at position 410 from gag (encoding amino acid E, glutamic acid) to cgg (amino acid R, arginine). The PCR product was digested with the appropriate enzymes, and isolated by agarose gel electrophoresis. The large fragment resulting from the BamHI/NheI digestion of pREFY2pref was also gel purified, and ligated to the PCR fragment above. Following transformation into E. coli, plasmid DNA was isolated and subsequently sequenced to confirm the presence of the E410R substitution. The amino acid sequence for Taq Δ271/F272M/F667Y/E410R DNA polymerase is shown at Figure 2.

Expression & Purification of the Taq Δ271/F272M/F667Y/E410R Polymerase

Vector pRE2 which carries the lambda p_L promoter was used with an *E. coli* strain which has the heat labile repressor protein cI857 to express the Taq Δ271/F272M/F667Y/E410R polymerase. This combination permits cultivation at 30°C followed by expression of a plasmid-borne protein at elevated temperatures such as 42°C. Liquid cultures were typically grown at 30°C to an OD_{600} of ~ 1.0, and then transferred to 42°C for ~ 2.5 hours. Bacterial cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris-HCl pH 8.5, 1 mM EDTA, 0.1% Tween-20, 0.1% Triton X-100, 10 mM $MgCl_2$, and 16 mM $(NH_4)_2SO_4$), and then heated at 80°C for 20 minutes to precipitate *E. coli* proteins. The heat lysate was clarified by centrifugation, and supplemented with 300 mM NaCl, and applied to a DE52 anion exchange column (commercially available from Whatman). The flow-through was diluted in Buffer A (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 10% glycerol, 0.1% Triton X-100 and 0.1% Tween-20) to reduce the NaCl concentration to 100 mM, and applied to a Heparin Sepharose column (commercially available from Pharmacia Inc.). The column was developed by linear gradient from 100 to 700 mM NaCl in Buffer A. The enzyme eluted at ~250mM NaCl. Fractions containing polymerase activity were pooled, concentrated on a Centriprep-50 apparatus (commercially available from Amicon) and dialyzed extensively against a final buffer containing 20 mM Tris-HCl pH 8.5, 50% glycerol, 0.1 mM EDTA, 0.5% Tween-20, 0.5% Triton X-100, 100 mM KCl and 1 mM DTT. The purity of the polymerase preparation was confirmed by SDS-PAGE.

Enzyme Characterization

1) Salt tolerance:

The Taq Δ271/F272M/F667Y/E410R DNA polymerase activity has been examined under a KCl titration experiment by using both activated salmon sperm DNA and primed M13 DNA as substrates. In both assays, Taq Δ271/F272M/F667Y E410R showed a decreased polymerase activity while increasing KCl concentration from 0 to 200 mM. However, the enzyme displays a much slower activity decrease compared to TS. Figure 4 plots the data from KCl titration of Taq Δ271/F272M/F667Y and Taq Δ271/F272M/F667Y/E410R using activated salmon sperm DNA as substrates. The 50% KCl inhibition for Taq Δ271/F272M/F667Y/E410R polymerase activity with activated salmon sperm or primed

M13 DNA are 120 mM and 100 mM, respectively compared to TS, which has a 50% KCl inhibition of 35 mM. The polymerase assay buffer contains: 25 mM TAPS (pH 9.3), 2 mM $MgCl_2$, 1 mM 2-mercaptoethanol and 200 mM each dNTP plus 0.05 Ci/mmol [α-^{33}P]-dATP. A comparison of salt tolerance data for Taq Δ271/F272M/F667Y and substitutions thereof is presented below in Table I.

TABLE I

Enzyme, substitution	Salt Tolerance
Taq Δ271/F272M/F667Y	35 mM
Taq Δ271/F272M/F667Y/E410R	135 mM
Taq Δ271/F272M/F667Y/E410M	125 mM
Taq Δ271/F272M/F667Y/E410W	125 mM
Taq Δ271/F272M/F667Y/E410H	110 mM

2) Thermostability at 95°C:

The thermostability of Taq Δ271/F272M/F667Y/E410R has been assayed as follows. First, the 95°C heating step was performed in a buffer containing 50 mM Tris-HCl pH 9.5, 5mM $MgCl_2$, 50μM each dNTP and 100ng M13 single strand DNA. Then 10 units of enzyme were mixed with the above solution and a time course performed by taking aliquots (20 μl each) and placing on ice. Next, dilutions were made in a buffer containing 10 mM Tris-HCl, pH 8.0, 1 mM 2-mercaptoethanal, 0.5% Tween-20, 0.5% Nonidet P-40. In the third step, the heated and diluted samples have been assayed for survivor polymerase activity under a standard polymerase assay condition described in section (1) but including 50 mM KCl. Figure 5 showed the thermostability assay of comparing Taq Δ271/F272M/F667Y/E410R with Amplitaq. The 50% inhibition time at 95°C for Taq Δ271/F272M/F667Y/E410R and Amplitaq are 25 min and 8min, respectively.

3) Processivity assay:

The processivity of Taq Δ271/F272M/F667Y/E410R has been examined in an enzyme dilution method, which insures that the polymerase activity is assayed for a single enzyme binding event. The assay buffer contains 15 mM Tris-HCl (pH 9.5), 3.5 mM $MgCl_2$, 100 mM each dNTP and 1μg P^{33} labeled primed M13. The primer extension experiment has been performed at 65°C for 90 seconds. The samples were analyzed on a 8% polyacrylamide-7 M urea sequencing gel. Taq Δ271/F272M/F667Y/E410R has an increased processivity of about 30 nucleotides per polymerase binding event. This is about a 7 to 8 fold increases compared to Taq Δ271/F272M/F667Y (4 nt/binding event).

4) Uniform termination events:

The new E to R amino acid modification discovered also results in increased uniformity in termination events during sequencing reactions containing net positive, negative, or neutrally charged dideoxynucleotide terminators. This results in an increased uniformity in electropherogram band intensity and an increase in the number of bases which can be basedcalled per sequence. For example, as shown in Figure 6, the average deviation of band intensity using Thermosequenase Version II is about a 30% deviation. However, as shown in Figure 7, a typical result using an E to R polymerase is about a 22% deviation. This improvement is significant. Portions of Figures 6 and 7 are magnified in Figures 8 through 10 for comparison purposes.

5) Ability to sequence difficult areas:

The new E to R amino acid modification discovered also results in an improved ability to sequence DNA's which contain "difficult to sequence" areas. Certain specific DNA sequences are extremely likely to cause sequencing DNA polymerases problems, resulting in a reduced quality of the sequence obtained (see Figure 11). Surprisingly, enzymes containing the E to R modification are much more likely to yield higher quality sequence data from DNA containing these difficult to sequence areas (see Figure 12).

EXAMPLE 2: TAQ D18A/E681R/F667Y POLYMERASE

We also constructed using standard techniques described above a full length version of Taq polymerase with the following substitutions : D18A/E681R/F667Y. In this enzyme, the D18A substitution removes the 5' to 3' exonuclease activity, rather than the deletion of

amino acids as in the Taq Δ271/F272M/F667Y DNA polymerase polypeptide. The E681R substitution is the position equivalent to E410R in Taq Δ271/F272M/F667Y DNA polymerase, and F667Y is the equivalent position to F396Y in Taq Δ271/F272M/F667Y DNA polymerase. This enzyme also has properties desirable for sequencing with dye terminators. The amino acid sequence of Taq D18A/E681R/F667Y DNA polymerase is shown at Figure 3.

Uniformity of positive terminator reactions is improved considerably with the substitutions at E681 as shown by the data in Table II below.

TABLE II

Enzyme, substitution	Uniformity (r.m.s.)
TSI, E681	0.52
TSI, E681R	0.39
TSI, E681H	0.37
TSI, E681I	0.4
TSI, E681M	0.31
TSI, E681W	0.34

Root mean square ("r.m.s.") is a measure of uniformity of a four color sequence reaction. This experiment used positive terminators (5 lysines in the linker) and standard sequencing reaction conditions. The improvement of 0.52 to below 0.45 shows a significant increase in uniformity for the sequencing reaction.

Figure 13 is a side-by-side comparison of electropherograms obtained from four color sequencing reactions conducted using D18A/F667Y DNA polymerases having various E681 substitutions as described at the left of each electropherogram. As shown in Figure 13, D18A/E681R/F667Y shows the most uniform peak heights and thus the most improvement in uniformity.

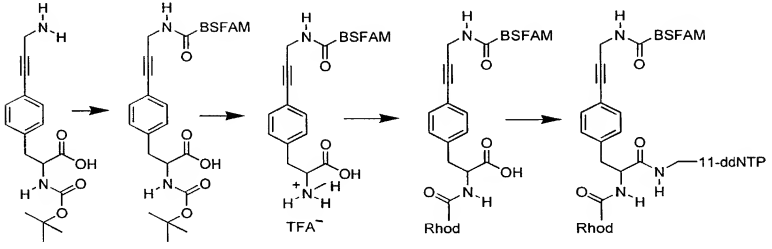
Figure 14 shows the relative reactivity compared to unlabelled ddNTPs evidenced in four color sequencing reactions which employed D18A/F667Y and various E681 substitutions therEof with various charged terminators.

Nucleic Acid Terminators

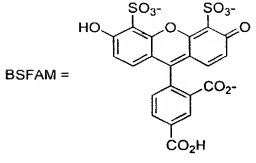
1. An example of charge modified reporters as applied to direct load

1.1 Chemistry

The following scheme was used to synthesize labeled ddNTPs with a charged reporter moiety. The linker was synthesized according to methods disclosed in U.S. Provisional Application No. 60/098,469 filed on August 31, 1998, the entire disclosure of which is hereby incorporated by reference herein.



1 2 3 4



Rhod = 5-R110, 5-ROX, 5-TAMRA, 5-REG

1.2 Discussion

4',5' Bis-sulfono-5-carboxyfluorescein (BSFAM) was attached to 4-propargylamino-N-α-*t*-butoxycarbonylphenylalanine by initial formation of the corresponding N-hydroxysuccinimide active ester using TSTU in DMF/diisopropylethylamine. Activation times were typically 15 minutes as observed by tlc before addition of the amino component. The product **1** was isolated by C18 RP-HPLC then treated with neat trifluoroacetic acid to remove the carbamate moiety, with the product **2** isolated by Et_2O precipitation. Attachment of the rhodamine moiety was carried out using 5-rhodamine hydroxysuccinimde active esters in DMSO/diisopropylethylamine. All the double dye cassettes were purified by reverse phase HPLC prior to conjugation to alkylamino ddNTPs using published methods (and as disclosed in U.S. Provisional Application No. 60/098,469 filed on August 31, 1998, the entire disclosure of which is hereby incorporated by reference herein). The labeled ddNTPs were purified by silica gel chromatography followed by ion exchange chromatography then reverse phase HPLC.

1.3 Experimental

All chemicals were purchased from Sigma, Aldrich, Fluka or Fisher Scientific unless stated otherwise. UV/visible spectra were recorded on a Perkin Elmer Lambda 20 UV/visible spectrophotometer in conjunction with Winlab™ software.

4-(propargylamido-4',5'-bissulfonatefluorescein)-N-α-*t*-butoxycarbonylphenylalanine (**1**)

4'-5'-bissulfono-5-carboxyfluorescein (100mg, 0.18mmol) was dissolved in DMF (4ml) then diisopropylethylamine (0.48ml, 15 eq.) and TSTU (65mg, 1.2eq.) added. The reaction mixture was stirred at room temperature for 1h. then 4-propargylamino-N-α-*t*-butoxycarbonylphenylalanine (69mg, 1.0eq) added. Stirring was continued for 3h. then the reaction mixture evaporated to dryness *in vacuo*. The product was isolated by reverse phase HPLC (C18, DeltaPak 15μ, 100A, 50x300μm) eluting with 0-100% eluant B over 60 min (A = 0.1M TEAB, B = 50% MeCN/0.1MTEAB v/v, 100ml/min.). The product (retention time 37 min.) was evaporated to dryness *in vacuo* then coevaporated with MeOH (3x10ml) before

lyophilization (100mg, 65%). UV/vis (1M triethylammonium bicarbonate pH 8.8) 495nm (24670), 465nm (shoulder, 9634), 312nm (6708).

4-(propargylamido-4',5'-bissulfonatefluorescein)-phenylalanine-α-ammonium trifluoroacetate **(2)**

4-(propargylamido-4',5'-bissulfonatefluorescein)-N-α-*t*-butoxycarbonylphenylalanine (100mg, 0.12mmol) was treated with trifluoroacetic acid (10ml) for 15min. then evaporated to dryness *in vacuo*. The residue was coevaporated with toluene (3x10ml) then the product precipitated by the addition of Et_2O (50ml). The solid formed was collected by filtration, washed with cold Et_2O (3x50ml) then dried under high vacuum (100mg, 99%). Rf (tlc, iPrOH:NH_4OH:H_2O (6:3:1)=0.

General methodology for the attachment of rhodamine dyes to **2 (3)**

4-(propargylamido-4',5'-bissulfonatefluorescein)-phenylalanine-α-ammonium trifluoroacetate **2** (0.1mmol) was dissolved in DMSO (1ml) then diisopropylethylamine (0.26ml, 15 eq.) and rhodamine-NHS active ester (1.5 eq.) added. The reaction mixture was stirred at room temperature for 16h, then evaporated to dryness *in vacuo*. The R110 analog was treated with triethyammonium bicarbonate solution (0.1M, 50ml) for 16h to remove the trifluoroacetimido protecting groups then the product purified by RP-HPLC using identical conditions to **1** unless stated. Retention times (BSFAM/R110 = 31min, BSFAM/R110 = 55min 0-100% B over 90 min, 100 ml/min., BSFAM/REG 54min 0-100%B over 90 min., 100ml/min, BSFAM/TAMRA = 52min 0-100% B over 90 min). All absorption spectra show the presence of both dyes.

General Methodology for Attachment of **3** to alkylamino-2',3'-dideoxynucleotide triphosphates **(4)**.

The double dye cassette (1mmol) was dissolved in DMF (5ml) then disuccinimdyl carbonate (4eq.) and DMAP (4eq.) were added at –60°C. The reaction mixture was stirred at –30°C for 15 min. then a solution of aminoalkyl-ddNTP (0.67eq., Na_2CO_3/$NaHCO_3$ pH 8.5) added.

The reaction was stirred at room temperature for 1h. then applied directly to a SiO_2 gel column. The product was eluted with iPrOH:NH_4OH:H_2O (4:5:1 v:v:v) then evaporated to dryness *in vacuo* before subsequent purification by ion exchange chromatography then C18 reverse phase HPLC as for **1**. Absorption spectra of each compound showed the presence of both dyes.

1.4 Comparative Electropherograms

One of the terminators (structure 4, Rhodamine =5-ROX and N = C) formed above was used in a sequencing reaction and run on a slab gel. The resulting electropherogram is shown in Figure 16 which provides an example of the increase in migration rate relative to sequence products of unincorporated bis-sulfonated fluorescein energy transfer terminators (and thermal breakdown products thereof) compared to the migration rate of the regular ET terminators.

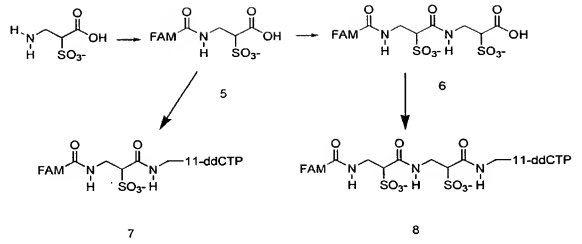
2. An example of a negatively charged linker arm as applied to direct load

2.1 Background

By incorporation of a number of charged amino acids onto a fluorescent reporter, it is possible to synthesize a labeled ddNTP containing extra negative charge that alters the mobility of the degradative by-products observed in a sequencing reaction.

2.2 CHEMISTRY

In order to determine the amount of negative charge required to remove colored by-products from the sequence ladder, fluorescein was attached to α-sulfo-β-alanine to form **5**. Compound **5** was attached to a 11-ddCTP (11=number of atoms in linker arm) to form **7**. A portion of **5** was attached to a second α-sulfo-β-alanine moiety to form **6** which was subsequently attached to 11-ddCTP to form **8**. A control ddNTP containing regular FAM attached to 11-ddCTP was also synthesized. The structures were run in a single color sequencing reaction to determine the effect of the charge on mobility.



As fluorescein carries a net 1- charge, compound 7 is considered as overall 2- linker arm, compound 8 has an overall 3- linker arm charge.

2.3 Experimental

N-5-carboxamidofluorescein-α-sulfo-β-alanine (5)

α-sulfo-β-alanine (59mg, 0.35mmol) was dissolved in DMF (2ml) then diisopropylethylamine (0.9mol, 15eq) added followed by 5-FAM-NHS active ester (200mg, 1.2eq.). The reaction mixture was stirred at room temperature for 3h. then evaporated to dryness *in vacuo*. The residue was coevaporated with MeOH (10ml) then the product isolated by C18 RP HPLC (A=0.1MTEAB, B=0.1MTEAB, 50%MeCN v/v) eluting with 0-100%B over 90 min at 100ml/min. ^{1}H nmr (300MHz, CD_3OD); 1.27(t, 24H, J=8.4Hz, $NCH_2\underline{CH_3}$), 3.05(q, 16H, J=8.4Hz, $N\underline{CH_2}CH_3$), 3.95-4.05(m, 3H, CH_2+$\underline{CH}SO_3$), 6.58(m, 3H, Ar-H), 6.85(d, 2H, J=11.0Hz, Ar-H), 7.30(d, 2H, J=11.0Hz, Ar-H), 8.02(d, 1H, J=7.6Hz, ArH), 8.45(s,1H,Ar-H).

N-(N-5-carboxamidofluorescein-α-sulfo-β-alanine)amido-α-sulfo-β-alanine (6)

N-5-carboxamidofluorescein-α-sulfo-β-alanine (5, 50mg, 0.095mmol) was dissolved in DMF (3ml) then diisopropylethylamine (0.25ml, 15eq.) and TSTU (42mg, 1.5eq.) added. The reaction mixture was stirred at room temperature for 1h. then α-sulfo-β-alanine (24mg, 1.5eq.) added. Stirring was continued for 3h. then the reaction evaporated to dryness *in vacuo*. The product was isolated by ion exchange chromatography (mono-Q column, A=0.1M TEAB, 40%MeCN v/v, B=1.0M TEAB, 40%MeCN v/v, 0-50%B over 22min., 50-75%B from 22-50min. 75-100%B from 50-70 min., 4ml/min., retention time = 75-80min.) then C18 RP HPLC (A=0.1M TEABB=0.1M TEAB/MeCN 50% v/v, 0-100%B over 90 min., 100ml/min, retention time = 33min.). Rf_1(PrOH6:ammonia3:water1v/v/v) 0.34.

General Methodology for Attachment of modified dyes to alkylamino-2',3'-dideoxynucleotide triphosphates (7,8).

The modified dye (1mmol) was dissolved in DMF (5ml) then disuccinimdyl carbonate (4eq.) and DMAP (4eq.) were added at –60°C. The reaction mixture was stirred at –30°C for 15 min. then a solution of aminoalkyl-ddNTP (0.67eq., Na_2CO_3/$NaHCO_3$ pH 8.5) added. The reaction was stirred at room temperature for 1h. then applied directly to a SiO_2 gel column. The product was eluted with iPrOH:NH_4OH:H_2O (4:5:1 v:v:v) then evaporated to dryness *in vacuo* before subsequent purification by ion exchange chromatography then C18 reverse phase HPLC as for **1**.

2.4 Results

Each labeled ddNTP was dissolved in sequencing buffer and subjected to several rounds of thermocycling. The products were separated on a sequencing gel and the electropherograms shown in Figure 17. Interpretation of the electropherogram provided the conclusion an overall 3- charge (i.e., structure 8) removed the colored by-products from the area of the electropherogram where true sequencing data would be obtained.

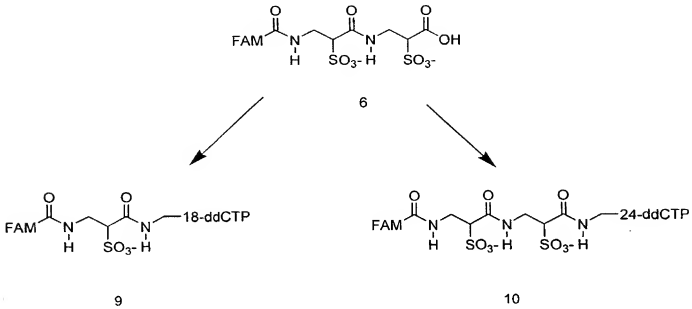
Figure 17 illustrates how the net negative charge of the dye labeled dideoxynucleotides affects their (and thermal breakdown products thereof) migration rate. As the net negative charge of the terminator increases, the migration rates of the various peaks seen (each of the peaks seen are either dye labeled dideoxynucleotides or thermal breakdown products thereof) increases (Figure 17). At an overall 3- charge (2- from linker, 1-from fluorescein) peaks are absent from the region of the electropherogram where true sequence data would normally be obtained.

3. Negatively charged extended linker arms as applied to direct load

3.1 Background

In order to improve the efficiency of incorporation of the modified terminator, a labeled terminator with a 3- charge on the linker arm was synthesized, this time containing an extended linker arm of 18 and 24 atoms.

3.2 Chemistry



3.3 Experimental

Compound 6, was attached to 18-ddCTP and 24-ddCTP using the standard protocol for attachment of labels to ddNTPs outlined in section 2.3. The method of purification was the same for 9 and 10.

Retention time of 9: Mono-Q™ ion exchange (47min)
Retention time of 10: Mono-Q™ ion exchange (42min)
C18 RP-HPLC (15min)

3.4 Sequencing Results

From the sequencing experiments it was clear that increasing the linker arm length improved incorporation of the terminator. This information, coupled to the presence of the 3- charge in the dye-linker structure led us to investigate rhodamine dyes with a 3- charged linker. This would permit four color sequencing.

As shown in Figure 18, it is possible to directly load a sequencing reaction with no clean-up procedure. Figure 18 shows no peaks resulting from unincorporated dye-labeled terminator in the sequence, thus demonstrating the utility of negatively charged terminators with respect to direct load sequencing.

3.5 Rhodamine Labeled Terminators Containing a 3- Linker Arm

The following chemistry was attempted to synthesize a set of four differently labeled terminators:

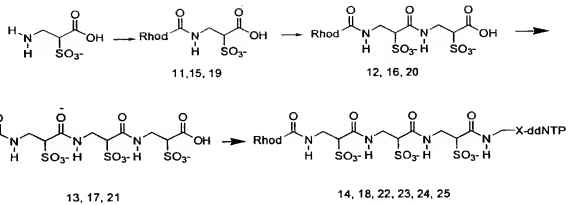


TABLE III

Compound Nos.	Rhod	X	N
11-14	REG	24	U
15-18	TAMRA	24	A
19-22	ROX	24	G
23	TAMRA	12	A
24	ROX	12	G
25	ROX	18	G

Rhod = rhodamine label, X = length of linker arm, N=base

3.6 Experimental

Compounds 11, 15, 19 were synthesized according to the method outlined for 5.
Compounds 12, 13,16,17,17,21 according to the method outlined for 6.
Compounds 14, 18,22-25 according to the general methodology for attachment of modified dyes to alkylamino-2',3'-dideoxynucleotide triphosphates (7,8).

3.7 Results and Discussion

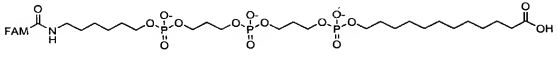
The labeled triphosphates 14, 18, 22 were used in a direct load sequencing experiment. Compound 14 in a direct load experiment showed no breakdown products and with TSII and

TaqERDAFY. Compounds 18 and 22 gave very dark sequencing bands and were observed to be forming an unexpected aggregate (as observed in the emission spectrum). The compounds also produced large colored blobs on a sequencing gel which interfered with interpretation of the sequence.

In order to overcome the aggregation effect, structures 23-25 were synthesized to investigate the effect of a shorter linker arm. Compound 23 has been shown to yield a clean sequence, 24 and 25 are awaiting testing. Structures 23-25 all have the expected rhodamine emission spectrum hence it appears that the aggregation problem may have been overcome.

4 Other examples of negatively charged linker arms

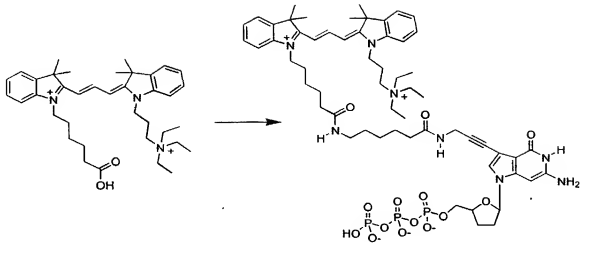
Other negatively charged linker arms have been synthesized and studied for example the phopshodiester structure shown below. The product was synthesized using phopshoramidite chemistry however it could also be synthesized via H-phosphonates, phosphoroimidazolides, or phosphotriester chemistry.



5. Examples of Terminators with Positively Charged Reporters

5.1 Background

In order to study positively charged structures, the following labeled terminator was synthesized.



26 27

5.2 Experimental

Compound 26 (10mg, 0.0134mmol) was dissolved in DMF (1ml) then diisopropylethylamine (23µl, 10eq.) added followed by PyBOP (14mg, 2.0eq.). The reaction mixture was stirred at room temperature for 15min. then a solution of 11-ddGTP (0.0083mmol, Na_2CO_3-$NaHCO_3$ pH 8.5) added in one portion. The reaction mixture was stirred at room temperature for 3h. then applied directly to a silica gel column. The product was eluted with iPrOH:NH_4OH:H_2O (6:3:1 v/v/v) then purified by ion exchange chromatography (as for 6) followed by C18 RP-HPLC (1.75µmol yield, 21%).

5.3 Sequencing Results

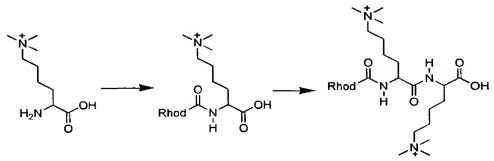
The electropherogram shown in Figure 19 was obtained when 27 was used in a sequencing reaction. The +2 charged terminator was used in a sequencing reaction and loaded directly on to a slab gel. The same experiment was repeated, however the reaction mixture was treated with phosphatase prior to loading on a gel to remove phosphates from the unincorporated dye-labeled dideoxynucleotides remaining in the reaction mixture. This leaves all terminator derived products with an overall positive charge causing them to migrate in the opposite direction as the sequence products during electrophoresis. It is clear from the

electropherogram that the colored by-products are absent from the sequence when phosphatase is used to break down the terminator products.

6. Positively charged extended linker arms as applied to direct load

6.1 Chemistry

Another example of dyes attached to a positively charged linker arm is shown below;



28 29

In this example, the rhodamine dye R6G is attached to ε-N,N,N-trimethyllysine which contains a formalized positive charge from the ε quaternary amine. The product (28) can be further modified to yield a +2 linker arm (29) by reaction with a further molecule of the charged amino acid. Further reaction(s) would generate the desired charged structure.

6.2 Experimental

α–N–(5–carboxamidorhodamine6G)-ε-N,N,N-trimethyllysine (28)

ε-N,N,N-trimethyllysine (68mg, 30.0mmol) was dissolved in DMF (6ml) then diisopropylethylamine (0.5ml, 10eq.) added followed by R6G-NHS active ester (200mg, 1.2eq.). The reaction mixture was stirred at room temperature for 16h. then evaporated to dryness *in vacuo*. The product was isolated by C18 RP HPLC (A=0.1M TEAB, B=0.1MTEAB/50%MeCN, 0-100%B over 50 min., 100ml/min). Retention time = 44min.

α-(α'-N-(5-carboxamidorhodamine6G)-ε'-N,N,N-trimethyllysine)-ε-N,N,N-trimethyllysine (29)

α-N-(5-carboxamidorhodamine6G)-ε-N,N,N-trimethyllysine 28 (100mg, 0.15mmol) was dissolved in DMF (5ml) then diisopropylethylamine (0.3ml, 15eq.) and TSTU (67mg, 1.5eq.) added. The reaction mixture was stirred at room temperature for 1h. then ε-N,N,N-trimethyllysine (50mg, 1.5eq.) added. The solution was stirred for a further 3h. then the reaction mixture was evaporated to dryness *in vacuo*. The product was isolated by C18 RP HPLC (A=0.1M TEAB, B=0.1MTEAB/50%MeCN, 0-100%B over 90 min., 100ml/min). Retention time = 60min.

TABLE IV

Abbreviations

Abbreviation	Definition
ddNTP	2'-3'-dideoxynucleoside triphosphate
ET	Energy Transfer
TSTU	2-Succinimido-1,1,3,3-tetramethyluronium tetrafluoroborate
PyBOP	Benzotrialzol-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate
DMF	N,N-dimethylformamide
RP HPLC	Reverse Phase High Performance Liquid Chromatography
Et_2O	Diethyl ether
DMSO	Dimethyl sulfoxide
TEAB	Triethylammonium bicarbonate
MeCN	Acetonitrile
iPrOH	Isopropanol
NH_4OH	Ammonium Hydroxide
BSFAM	4',5' Bis-sulfono-5-carboxyfluorescein
R110	Rhodamine 110
REG or R6G	Carboxyrhodamine6G
TAMRA	Tertamethylrhodamine
ROX	Carboxy-X-rhodamine
DMAP	4-dimethylaminopyridine
11-ddGTP	2',2'-dideoxyguanosine triphosphate with an 11 atom linker arm
NHS	N-hydroxysuccinimide

Although various embodiments of the instant invention are described in detail above, the instant invention is not limited to such specific examples. Various modifications will be readily apparent to one of ordinary skill in the art and fall within the spirit and scope of the following appended claims.

CLAIMS

What is claimed is:

1. A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 2.

2. A purified recombinant thermostable DNA polymerase comprising an amino acid substitution corresponding to the substitution E681R in the amino acid sequence listing of Figure 2.

3. An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence corresponding to the amino acid sequence set forth in Figure 2.

4. A recombinant DNA vector that comprises the nucleic acid of Claim 3.

5. A recombinant host cell transformed with the vector of Claim 4.

6. The recombinant host cell of Claim 5 that is *E. coli.*

7. A method of sequencing DNA comprising the step of generating chain terminated fragments from the DNA template to be sequenced with the DNA polymerase of Claim 1 in the presence of at least one chain terminating agent and one or more nucleotide triphosphates, and determining the sequence of said DNA from the sizes of said fragments.

8. A method according to Claim 7, wherein the chain terminating agent comprises a labeled nucleic acid terminator having a net positive or a net negative charge.

9. A method for synthesizing a fluorescently labeled polynucleotide, said method comprising the step of mixing a DNA polymerase according to Claim 1 with a primed template.

10. A method according to Claim 9, wherein the primed template is a primed template in a chain termination sequencing reaction.

11. A method according to Claim 9, wherein the primed template is a primed template in a polymerase chain reaction.

12. A kit for fluorescently labeling a polynucleotide, the kit comprising a DNA polymerase according to Claim 1 and a fluorescently labeled nucleotide.

13. A kit according to Claim 12, wherein the fluorescently labeled nucleotide comprises a nucleic acid terminator having a net negative or a net positive charge.

14. A kit for sequencing DNA comprising the DNA polymerase of Claim 1.

15. A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 3.

16. A DNA polymerase comprising an amino acid substitution corresponding to the substitution E681R in the amino acid sequence listing of Figure 3.

17. An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence corresponding to the amino acid sequence set forth in Figure 3.

18. A recombinant DNA vector that comprises the nucleic acid of Claim 17.

19. A recombinant host cell transformed with the vector of Claim 18.

20. The recombinant host cell of Claim 18 that is *E. coli.*

21. A method of sequencing DNA comprising the step of generating chain terminated fragments from the DNA template to be sequenced with the DNA polymerase of Claim 16 in the presence of at least one chain terminating agent and one or more nucleotide triphosphates, and determining the sequence of said DNA from the sizes of said fragments.

22. A method according to Claim 21, wherein the chain terminating agent comprises a labeled nucleic acid terminator having a net positive or a net negative charge.

23. A method for synthesizing a fluorescently labeled polynucleotide, said method comprising the step of mixing a DNA polymerase according to Claim 16 with a primed template.

24. A method according to Claim 23, wherein the primed template is a primed template in a chain termination sequencing reaction.

25. A method according to claim 23, wherein the primed template is a primed template in a polymerase chain reaction.

26. A kit for fluorescently labeling a polynucleotide, the kit comprising a DNA polymerase according to Claim 16 and a fluorescently labeled nucleotide.

27. A kit according to Claim 26, wherein the fluorescently labeled nucleotide comprises a nucleic acid terminator having a net negative or a net positive charge.

28. A kit for sequencing DNA comprising the DNA polymerase of Claim 16.

(12) INTERNATIONAL PUBLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 March 2001 (01.03.2001)

PCT

(10) International Publication Number
WO 01/14568 A1

(51) International Patent Classification[7]: C12N 15/54. 15/70, 9/12, 1/21, C12Q 1/68, C12P 19/34 // (C12N 1/21, C12R 1:19)

(21) International Application Number: PCT/US00/22150

(22) International Filing Date: 10 August 2000 (10.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/150,167 21 August 1999 (21.08.1999) US
60/154,739 17 September 1999 (17.09.1999) US

(71) Applicant *(for all designated States except US)*: AMERSHAM PHARMACIA BIOTECH INC. [US/US]; 800 Centennial Avenue, P.O. Box 1327, Piscataway, NJ 08855-1327 (US).

(72) Inventors; and

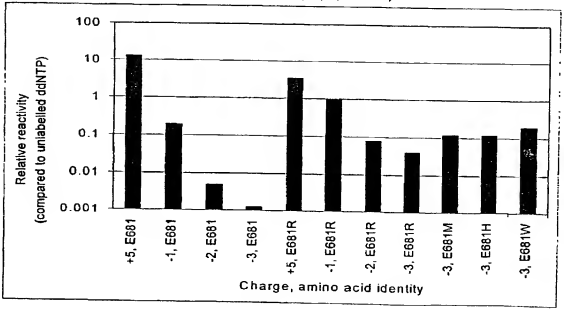
(75) Inventors/Applicants *(for US only)*: DAVIS, Maria [US/US]; P.O. Box 1327, 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US). NELSON, John [US/US]; P.O. Box 1327, 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US). KUMAR, Shiv [US/US]; P.O. Box 1327, 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US). FINN, Patrick, J. [GB/US]; P.O. Box 1327, 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US). NAMPALLI, Satyaam [IN/US]; P.O. Box 1327, 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US). FLICKE, Parke [US/US]; P.O. Box 1327, 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US).

(74) Agent: MALIA, Victoria, M.; Amersham Pharmacia Biotech Inc., P.O. Box 1327, 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,

[Continued on next page]

(54) Title: *TAQ* DNA POLYMERASES HAVING AN AMINO ACID SUBSTITUTION AT E681 AND HOMOLOGS THEREOF EXHIBITING IMPROVED SALT TOLERANCE



(57) Abstract: Thermostable DNA polymerases having an E410R substitution which result in a substantial improvement of signal uniformity compared to Taq Δ271/F272M/F667Y DNA polymerase. The instant DNA polymerases possess improved salt tolerance and have been shown to modulate the incorporation of terminators having a net positive or a net negative charge during the sequencing reaction.

WO 01/14568 A1

Taq DNA polymerase:

MRGMLPLFEPKGRVLLVDGHHLAYRTFHALKGLTTSRGEPVQAVYGF
AKSLLKALKEDGDAVIVVFDAKAPSFRHEAYGGYKAGRAPTPEDFPRQ
LALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADK
DLYQLLSDRIHVLHPEGYLITPAWLWEKYGLRPDQWADYRALTGDES
DNLPGVKGIGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDD
LKLSWDLAKVRTDLPLEVDFAKRREPDRERLRAFLERLEFGSLLHEFGL
LESPKALEEAPWPPPEGAFVGFVLSRKEPMWADLLALAAARGGRVHR
APEPYKALRDLKEARGLLAKDLSVLALREGLGLPPGDDPMLLAYLLDP
SNTTPEGVARRYGGEWTEEAGERAALSERLFANLWGRLEGEERLLWL
YREVERPLSAVLAHMEATGVRLDVAYLRALSLEVAEEIARLEAEVFRL
AGHPFNLNSRDQLERVLFDELGLPAIGKTEKTGKRSTSAAVLEALREAH
PIVEKILQYRELTKLKSTYIDPLPDLIHPRTGRLHTRFNQTATATGRLSSS
DPNLQNIPVRTPLGQRIRRAFIAEEGWLLVALDYSQIELRVLAHLSGDE
NLIRVFQEGRDIHTETASWMFGVPREAVDPLMRRAAKTINFGVLYGMS
AHRLSQELAIPYEEAQAFIERYFQSFPKVRAWIEKTLEEGRRRGYVETLF
GRRRYVPDLEARVKSVREAAERMAFNMPVQGTAADLMKLAMVKLFP
RLEEMGARMLLQVHDELVLEAPKERAEAVARLAKEVMEGVYPLAVPL
EVEVGIGEDWLSAKE

Fig. 1

Taq Δ271/F272M/F667Y/**E681R** DNA polymerase:

MLERLEFGSLLHEFGLLESPKALEEAPWPPPEGAFVGFVLSRKEPMWA
DLLALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLSVLALREGLG
LPPGDDPMLLAYLLDPSNTTPEGVARRYGGEWTEEAGERAALSERLFA
NLWGRLEGEERLLWLYREVERPLSAVLAHMEATGVRLDVAYLRALSL
EVAEEIARLEAEVFRLAGHPFNLNSRDQLERVLFDELGLPAIGKTEKTG
KRSTSAAVLEALREAHPIVEKILQYRELTKLKSTYIDPLPDLIHPRTGRL
HTRFNQTATATGRLSSSDPNLQNIPVRTPLGQRIRRAFIAEEGWLLVAL
DYSQIELRVLAHLSGDENLIRVFQEGRDIHTETASWMFGVPREAVDPL
MRRAAKTINYGVLYGMSAHRLSQRLAIPYEEAQAFIERYFQSFPKVRA
WIEKTLEEGRRRGYVETLFGRRRYVPDLEARVKSVREAAERMAFNMP
VQGTAADLMKLAMVKLFPRLEEMGARMLLQVHDELVLEAPKERAEA
VARLAKEVMEGVYPLAVPLEVEVGIGEDWLSAKE

Fig. 2

Taq D18A/E681R/F667Y DNA polymerase:

MRGMLPLFEPKGRVLLVAGHHLAYRTFHALKGLTTSRGEPVQAVYGF
AKSLLKALKEDGDAVIVVFDAKAPSFRHEAYGGYKAGRAPTPEDFPRQ
LALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADK
DLYQLLSDRIHVLHPEGYLITPAWLWEKYGLRPDQWADYRALTGDES
DNLPGVKGIGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDD
LKLSWDLAKVRTDLPLEVDFAKRREPDRERLRAFLERLEFGSLLHEFGL
LESPKALEEAPWPPPEGAFVGFVLSRKEPMWADLLALAAARGGRVHR
APEPYKALRDLKEARGLLAKDLSVLALREGLGLPPGDDPMLLAYLLDP
SNTTPEGVARRYGGEWTEEAGERAALSERLFANLWGRLEGEERLLWL
YREVERPLSAVLAHMEATGVRLDVAYLRALSLEVAEEIARLEAEVFRL
AGHPFNLNSRDQLERVLFDELGLPAIGKTEKTGKRSTSAAVLEALREAH
PIVEKILQYRELTKLKSTYIDPLPDLIHPRTGRLHTRFNQTATATGRLSSS
DPNLQNIPVRTPLGQRIRRAFIAEEGWLLVALDYSQIELRVLAHLSGDE
NLIRVFQEGRDIHTETASWMFGVPREAVDPLMRRAAKTINYGVLYGMS
AHRLSQRLAIPYEEAQAFIERYFQSFPKVRAWIEKTLEEGRRRGYVETL
FGRRRYVPDLEARVKSVREAAERMAFNMPVQGTAADLMKLAMVKLF
PRLEEMGARMLLQVHDELVLEAPKERAEAVARLAKEVMEGVYPLAVP
LEVEVGIGEDWLSAKE

Fig. 3

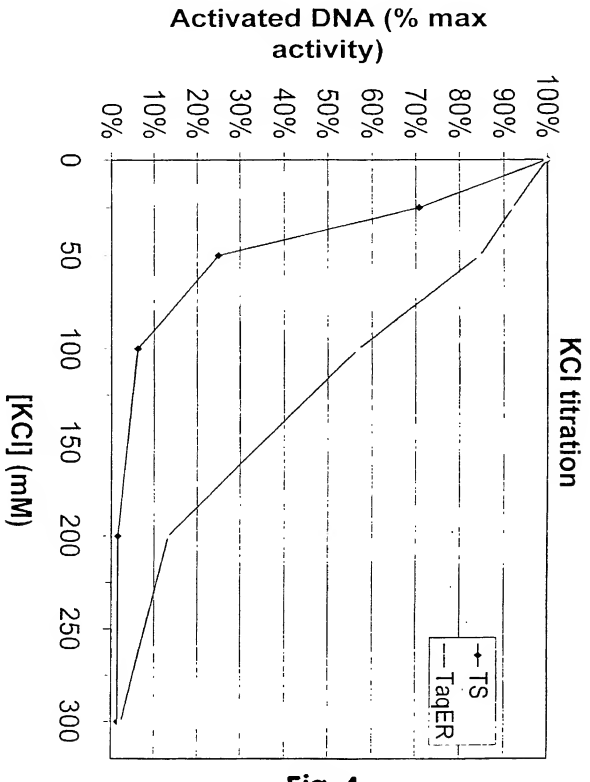

Activated DNA (% max activity)
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
0
50
100
150
200
250
300
[KCl] (mM)
KCl titration
TS
TaqER

Fig. 4

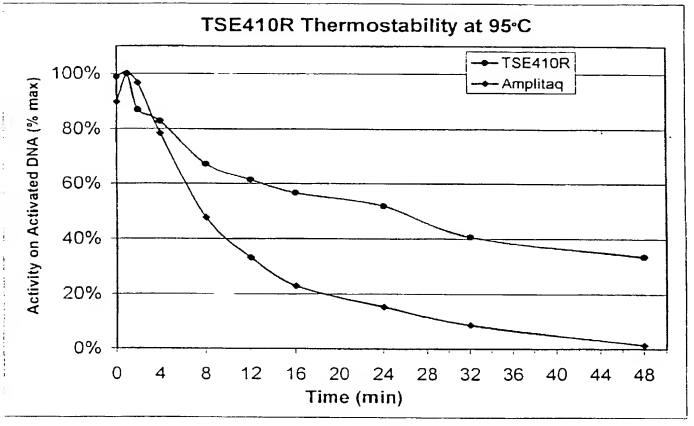

TSE410R Thermostability at 95°C
TSE410R
Amplitaq
Activity on Activated DNA (% max)
100%
80%
60%
40%
20%
0%
0
4
8
12
16
20
24
28
32
36
40
44
48
Time (min)

Fig. 5

Fig. 6

Fig. 7

Fig. 8

Difficult template, TSII ER

Fig. 9

Fig. 10a

DNA sequence using TSII.
Figure shows example of a strong \ relative to following \'s.

Fig. 10b

Same sequence using an E to R Polymerase. The strong \ relative to following \'s is eliminated.

Fig. 11a

DNA sequence using TSII. Figure shows multiple examples of weak I 's after preceding I 's.

Fig. 11b

Same sequence using an E to R Polymerase. The weak I after a preceding I is eliminated.

Fig. 12a

DNA sequence using TSII. Figure shows example of a weak G after preceding \.

Fig. 12b

Same sequence using an E to R Polymerase. The weak G after preceding \ is eliminated.

681
E
H
R
M
W
AGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGT
AGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTC
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGT
AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG
AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG

Fig. 13

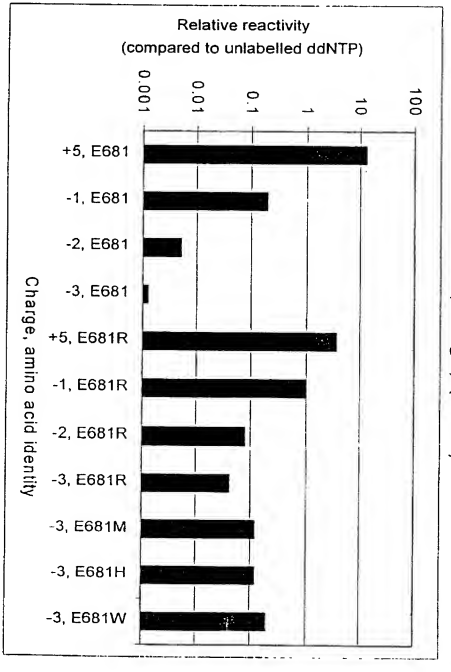

Relative reactivity
(compared to unlabelled ddNTP)
0.001
0.01
0.1
1
10
100
+5, E681
-1, E681
-2, E681
-3, E681
+5, E681R
-1, E681R
-2, E681R
-3, E681R
-3, E681M
-3, E681H
-3, E681W
Charge, amino acid identity
fluorescein-(charge)-(linker)-ddCTP

Fig. 14

15/19

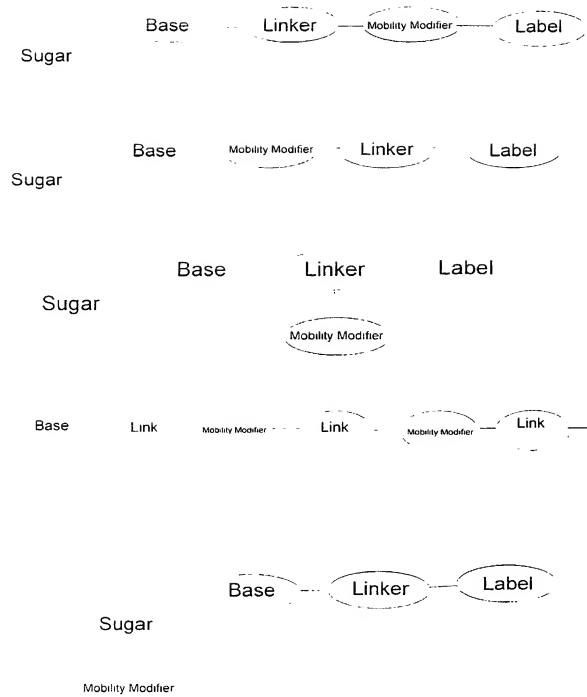

Base
Linker
Mobility Modifier
Label
Sugar
Base
Mobility Modifier
Linker
Label
Sugar
Base
Linker
Label
Sugar
Mobility Modifier
Base
Link
Mobility Modifier
Link
Mobility Modifier
Link
Label
Sugar
Base
Linker
Label
Sugar
Mobility Modifier

Fig. 15

Fig. 16a

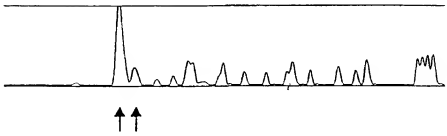
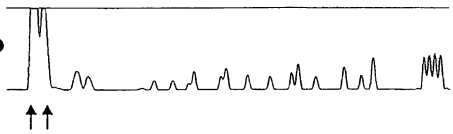


Fig. 16b



Faster Slower

Comparison of Regular v. Bis-sulfonated Fluorescein ET Terminators

Fig. 17a

Fig. 17b

Fig. 17c

Net -3 charge terminator (10) reaction, directly loaded

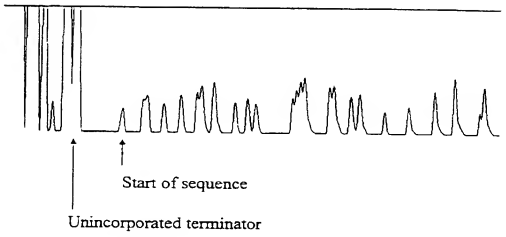


Fig. 18

Fig. 19a

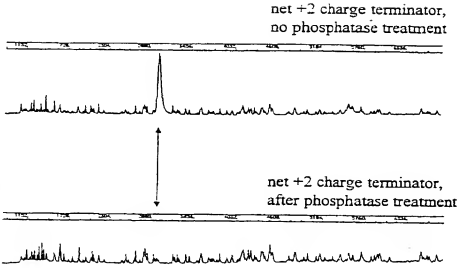


Fig. 19b

Docket No.: PB9944
Application No.: 10/049,358
Filing Date: to be assigned
Group Art Unit: to be assigned
Examiner: to be assigned
Declaration Submitted After Initial Filing

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

***TAQ* DNA Polymerases Having an Amino Acid Substitution at E681 and Homologs Thereof Exhibiting Improved Salt Tolerance**

the specification of which

[ ] is attached hereto.
OR
[X] was filed on **August 10, 2000** as United States Application No. or PCT International Application No. **PCT/US00/22150** and was amended on _____.(if applicable)

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional patent application(s) listed below:

60/148,012	**August 10, 1999**
(Application Serial No.)	(Filing Date)

I hereby claim the benefit under 35 U.S.C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application(s) designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, CFR Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

PCT/US00/22150	**August 10, 2000**
(Application Serial No.)	(Filing Date)

As a named inventor, I hereby appoint the following attorneys or agents to prosecute this application and transact all business in the United States Patent and Trademark Office connected therewith:

Customer Number: 22840

Royal N. Ronning, Jr. (Reg. No. 32,529)

Stephen G. Ryan (Reg. No. 39,015)

Robert F. Chisholm (Reg. No. 39,939)

Send correspondence to: Amersham Biosciences Corp
800 Centennial Avenue
Piscataway, New Jersey 08855

Direct telephone calls to: (732) 457-8423

Direct facsimiles to: (732) 457-8463

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-US

Full name of first inventor: Maria Davis

Inventor's signature: Maria Davis

Date: March 15, 2002

Post Office Address: 800 Centennial Avenue NJ
Piscataway, New Jersey 08855 US

Citizenship: United States

2-US

Full name of second inventor: John Nelson

Inventor's signature: John Nelson

Date: March, 19, 2002

Post Office Address: 800 Centennial Avenue NJ
Piscataway, New Jersey 08855 US

Citizenship: United States

3-US

Full name of third inventor: Shiv Kumar

Inventor's signature: Shiv Kumar

Date: March 15, 2002

Post Office Address: 800 Centennial Avenue NJ
Piscataway, New Jersey 08855 US

Citizenship: United States

4-00

Full name of fourth inventor: Patrick Finn

Inventor's signature:

Date: 15 MARCH 2002

Post Office Address: 800 Centennial Avenue NJ
Piscataway, New Jersey 08855 US

Citizenship: Great Britain

5-00

Full name of fifth inventor: Satyam Nampalli

Inventor's signature:

Date: 15 March 2002

Post Office Address: 800 Centennial Avenue NJ
Piscataway, New Jersey 08855 US

Citizenship: India

6-00

Full name of sixth inventor: Parke Flick

Inventor's signature:

Date: 3/15/02

Post Office Address: 800 Centennial Avenue NJ
Piscataway, New Jersey 08855 US

Citizenship: United States

SEQUENCE LISTING

<110> Davis, Maria
Nelson, John
Kumar, Shiv
Finn, Patrick J.
Nampalli, Satyam
Flick, Parke

<120> TAQ DNA Polymerase Having an Amino Acid Substitution at E681 and Homologs Thereof Exhibiting Improved Salt Tolerance

<130> PB9944

<140> PCT/US00/22150
<141> 2000-08-10

<150> 60/148,012
<151> 1999-08-10

<160> 3

<170> PatentIn Ver. 2.1

<210> 1
<211> 832
<212> PRT
<213> Thermus aquaticus

<400> 1

```
Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
  1               5                  10                  15

Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
             20                  25                  30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
         35                  40                  45

Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
     50                  55                  60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
 65                  70                  75                  80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
                 85                  90                  95
```

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
100 105 110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
115 120 125

Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
130 135 140

Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
145 150 155 160

Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
165 170 175

Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
180 185 190

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
195 200 205

Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
210 215 220

Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
225 230 235 240

Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
245 250 255

Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
260 265 270

Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
275 280 285

Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
290 295 300

Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
305 310 315 320

Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
325 330 335

Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu
340 345 350

```
Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro
        355                 360                 365

Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn
    370                 375                 380

Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu
385                 390                 395                 400

Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu
                405                 410                 415

Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu
            420                 425                 430

Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly
        435                 440                 445

Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala
    450                 455                 460

Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His
465                 470                 475                 480

Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp
                485                 490                 495

Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg
            500                 505                 510

Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile
        515                 520                 525

Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr
    530                 535                 540

Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu
545                 550                 555                 560

His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser
                565                 570                 575

Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln
            580                 585                 590

Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala
        595                 600                 605
```

```
Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly
    610                 615                 620

Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr
625                 630                 635                 640

Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro
                645                 650                 655

Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly
            660                 665                 670

Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu
        675                 680                 685

Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg
    690                 695                 700

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val
705                 710                 715                 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg
                725                 730                 735

Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
            740                 745                 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
        755                 760                 765

Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His
    770                 775                 780

Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala
785                 790                 795                 800

Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
                805                 810                 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
            820                 825                 830
```

<210> 2

<211> 560
<212> PRT
<213> Thermus aquaticus

<400> 2

```
Met Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 1               5                  10                  15

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
            20                  25                  30

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
        35                  40                  45

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
    50                  55                  60

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
65                  70                  75                  80

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
                85                  90                  95

Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
            100                 105                 110

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
        115                 120                 125

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
    130                 135                 140

Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
145                 150                 155                 160

Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
                165                 170                 175

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
            180                 185                 190

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
        195                 200                 205

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
    210                 215                 220

Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
```

```
225                 230                 235                 240

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
                245                 250                 255

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
            260                 265                 270

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
        275                 280                 285

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
    290                 295                 300

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
305                 310                 315                 320

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
                325                 330                 335

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
            340                 345                 350

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
        355                 360                 365

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
    370                 375                 380

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Tyr Gly Val Leu Tyr
385                 390                 395                 400

Gly Met Ser Ala His Arg Leu Ser Gln Arg Leu Ala Ile Pro Tyr Glu
                405                 410                 415

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
            420                 425                 430

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
        435                 440                 445

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
    450                 455                 460

Arg Val Lys Ser Val Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
465                 470                 475                 480

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
```

```
                485                 490                 495

Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His
            500                 505                 510

Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala
        515                 520                 525

Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
    530                 535                 540

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
545                 550                 555                 560
```

<210> 3
<211> 830
<212> PRT
<213> Thermus aquaticus

<400> 3

```
Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
  1               5                  10                  15

Val Ala Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
            20                  25                  30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
        35                  40                  45

Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
    50                  55                  60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
65                  70                  75                  80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
                85                  90                  95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
            100                 105                 110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
        115                 120                 125
```

```
Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
    130                 135                 140

Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
145                 150                 155                 160

Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
                165                 170                 175

Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
            180                 185                 190

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
        195                 200                 205

Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
    210                 215                 220

Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
225                 230                 235                 240

Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
                245                 250                 255

Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
            260                 265                 270

Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
        275                 280                 285

Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
    290                 295                 300

Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
305                 310                 315                 320

Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
                325                 330                 335

Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu Ala Lys
            340                 345                 350

Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro Pro Gly
        355                 360                 365

Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr Thr
    370                 375                 380
```

```
Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Glu Ala
385                 390                 395                 400

Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu Trp Gly
                405                 410                 415

Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu
            420                 425                 430

Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly Val Arg
        435                 440                 445

Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu
    450                 455                 460

Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe
465                 470                 475                 480

Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
                485                 490                 495

Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
            500                 505                 510

Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu
        515                 520                 525

Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile
    530                 535                 540

Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr
545                 550                 555                 560

Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
                565                 570                 575

Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
            580                 585                 590

Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp
        595                 600                 605

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
    610                 615                 620

Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr
625                 630                 635                 640
```

Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met
645 650 655

Arg Arg Ala Ala Lys Thr Ile Asn Tyr Gly Val Leu Tyr Gly Met Ser
660 665 670

Ala His Arg Leu Ser Gln Arg Leu Ala Ile Pro Tyr Glu Glu Ala Gln
675 680 685

Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
690 695 700

Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr
705 710 715 720

Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys
725 730 735

Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
740 745 750

Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
755 760 765

Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu
770 775 780

Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu
785 790 795 800

Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu
805 810 815

Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

SEQUENCE LISTING

<110> Davis, Maria
Nelson, John
Kumar, Shiv
Finn, Patrick
Nampalli, Satyam
Flick, Parke

<120> TAQ DNA Polymerase Having an Amino Acid Substitution at E681 and Homologs Thereof Exhibiting Improved Salt Tolerance

<130> PB9944

<140> 10/049,358
<141> To be assigned

<150> PCT/US00/22150
<151> 2000-08-10

<150> 60/148,012
<151> 1999-08-10

<160> 3

<170> PatentIn Ver. 2.1

<210> 1
<211> 832
<212> PRT
<213> Thermus aquaticus

<400> 1

```
Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
 1               5                  10                  15

Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
            20                  25                  30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
        35                  40                  45

Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
    50                  55                  60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
65                  70                  75                  80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
                85                  90                  95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
            100                 105                 110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
        115                 120                 125
```

```
Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
    130                 135                 140

Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
145                 150                 155                 160

Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
                165                 170                 175

Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
            180                 185                 190

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
        195                 200                 205

Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
    210                 215                 220

Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
225                 230                 235                 240

Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
                245                 250                 255

Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
            260                 265                 270

Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
        275                 280                 285

Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
    290                 295                 300

Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
305                 310                 315                 320

Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
                325                 330                 335

Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu
            340                 345                 350

Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro
        355                 360                 365

Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn
    370                 375                 380

Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu
385                 390                 395                 400

Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu
                405                 410                 415

Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu
            420                 425                 430
```

Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly
435 440 445

Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala
450 455 460

Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His
465 470 475 480

Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp
485 490 495

Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg
500 505 510

Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile
515 520 525

Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr
530 535 540

Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu
545 550 555 560

His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser
565 570 575

Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln
580 585 590

Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala
595 600 605

Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly
610 615 620

Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr
625 630 635 640

Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro
645 650 655

Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly
660 665 670

Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu
675 680 685

Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg
690 695 700

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val
705 710 715 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg
725 730 735

```
Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
            740                 745                 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
        755                 760                 765

Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His
    770                 775                 780

Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala
785                 790                 795                 800

Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
                805                 810                 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
            820                 825                 830
```

<210> 2
<211> 560
<212> PRT
<213> Thermus aquaticus

<400> 2

```
Met Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
  1               5                  10                  15

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
            20                  25                  30

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
        35                  40                  45

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
    50                  55                  60

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
65                  70                  75                  80

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
                85                  90                  95

Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
            100                 105                 110

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
        115                 120                 125

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
    130                 135                 140

Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
```

145 150 155 160

Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
165 170 175

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
180 185 190

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
195 200 205

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
210 215 220

Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
225 230 235 240

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
245 250 255

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
260 265 270

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
275 280 285

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
290 295 300

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
305 310 315 320

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
325 330 335

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
340 345 350

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
355 360 365

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
370 375 380

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Tyr Gly Val Leu Tyr
385 390 395 400

Gly Met Ser Ala His Arg Leu Ser Gln Arg Leu Ala Ile Pro Tyr Glu
405 410 415

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
420 425 430

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
435 440 445

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala

```
    450                 455                 460

Arg Val Lys Ser Val Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
465                 470                 475                 480

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
                485                 490                 495

Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His
            500                 505                 510

Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala
        515                 520                 525

Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
    530                 535                 540

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
545                 550                 555                 560
```

<210> 3
<211> 830
<212> PRT
<213> Thermus aquaticus

<400> 3

```
Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
  1               5                  10                  15

Val Ala Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
            20                  25                  30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
        35                  40                  45

Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
    50                  55                  60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
65                  70                  75                  80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
                85                  90                  95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
            100                 105                 110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
        115                 120                 125

Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
    130                 135                 140
```

```
Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
145                 150                 155                 160

Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
                165                 170                 175

Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
            180                 185                 190

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
        195                 200                 205

Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
    210                 215                 220

Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
225                 230                 235                 240

Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
                245                 250                 255

Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
            260                 265                 270

Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
        275                 280                 285

Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
    290                 295                 300

Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
305                 310                 315                 320

Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
                325                 330                 335

Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu Ala Lys
            340                 345                 350

Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro Pro Gly
        355                 360                 365

Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr Thr
    370                 375                 380

Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Glu Ala
385                 390                 395                 400

Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu Trp Gly
                405                 410                 415

Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu
            420                 425                 430

Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly Val Arg
        435                 440                 445
```

```
Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu
    450                 455                 460

Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe
465                 470                 475                 480

Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
                485                 490                 495

Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
            500                 505                 510

Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu
        515                 520                 525

Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile
    530                 535                 540

Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr
545                 550                 555                 560

Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
                565                 570                 575

Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
            580                 585                 590

Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp
        595                 600                 605

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
    610                 615                 620

Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr
625                 630                 635                 640

Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met
                645                 650                 655

Arg Arg Ala Ala Lys Thr Ile Asn Tyr Gly Val Leu Tyr Gly Met Ser
            660                 665                 670

Ala His Arg Leu Ser Gln Arg Leu Ala Ile Pro Tyr Glu Glu Ala Gln
        675                 680                 685

Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
    690                 695                 700

Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr
705                 710                 715                 720

Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys
                725                 730                 735

Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
            740                 745                 750
```

```
Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
        755                 760                 765

Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu
    770                 775                 780

Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu
785                 790                 795                 800

Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu
                805                 810                 815

Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
            820                 825                 830
```